Selected Weighted Adaptive Coding

Yoav Gross¹, Shmuel T. Klein², Elina Opalinsky¹ and Dana Shapira ¹

¹ Ariel University ² Bar Ilan University

1 Introduction

Weighted Coding

3 Selective Weighted Coding



Outline

Introduction

- Weighted Coding
- Selective Weighted Coding
- **4** Experimental Results

Selective Weighted Coding

Experimental Results

Data Compression

Input:

- Text
$$T = x_1 \cdots x_n$$

- Alphabet
$$\Sigma = \{\sigma_1, \dots, \sigma_m\}$$
, which occur $\{w_1, \dots, w_n\}$ times in T .

Goal:

Assign codewords with lengths $\{\ell_1, \ldots, \ell_n\}$ bits so that $\sum_{i=1}^n w_i \ell_i$ is *minimized*

Selective Weighted Coding

Experimental Results

Data Compression

Static:

The model - the distribution of the encoded elements

- Given in advance.
- Gathered in a first scan preprocess.

Adaptive:

- The model is learned incrementally.

Selective Weighted Coding

Experimental Results

Adaptive compression

Three main steps:

- **1** read the following symbol;
- **2 encode** *according to the current model*;
- **update** the model (increment the frequency of the currently read symbol).

Selective Weighted Coding

Experimental Results

Dynamic Huffman

 x_i is encoded with tree of $1, \ldots, x_{i-1}$.

Backward Looking:

- Only one pass.
- No need to transmit the tree.

Selective Weighted Coding

Experimental Results

Adaptive Coding

Backward Looking:

Base the current model on the past.

Forward Looking:

- Exact statistics.
- Base the current model on the future.

<ロト < 回 ト < 三 ト < 三 ト < 三 ト < 三 か Q () 8/34

Selective Weighted Coding

Experimental Results

Adaptive Coding

Backward Looking:

- "Selfish" behavior.
- Increments the frequency.

Forward Looking:

- "Altruistic" approach.
- Decrements the frequency.

Selective Weighted Coding

Experimental Results

Forward Coding vs. Static

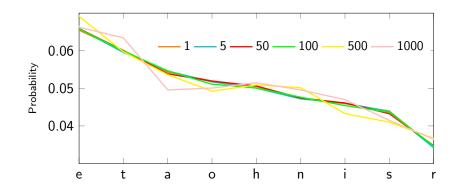
Theorem

For a given distribution of frequencies over an alphabet of size m, the encoded file by FORWARD is better than the encoded file by STATIC by at least m - 1 bits.

Selective Weighted Coding

Experimental Results

Distribution over a Subset of Elements



<ロト < 回 ト < 巨 ト < 巨 ト < 巨 ト 三 の Q () 11/34

7

Selective Weighted Coding

Experimental Results

Selective encoding/decoding

SELECTIVE-ENCODE($T = x_1 \cdots x_n$) / SELECTIVE-DECODE ($\mathcal{E}(T)$)

- 1 initialize the model
- 2 initialize a random bit generator
- 3 for $i \leftarrow 1$ to n do
- 4 encode / decode x_i according to the current model
- 5 $bit \leftarrow random()$
- 6 if bit = 1 then
 - Update the model

Selective Weighted Coding

Outline

Introduction

Weighted Coding

Selective Weighted Coding

4 Experimental Results

<ロト < 回 ト < 巨 ト < 巨 ト < 巨 ト 三 の < C 13/34

Intuition:

- Static: Fixed codewords, governed solely by the frequency.
- Adaptive: Based on the currently known statistics an estimate.
- Forward: Adaptive encoding based on the true frequencies.

All positions in the input file are treated equally.

- Weighted: Closer positions get higher priority

Selective Weighted Coding

Experimental Results

Backward Weighted Coding

$W(g, \sigma, i)$ definition:

Given T = T[1, n] over Σ , define a general weight $W(g, \sigma, i)$ - $g : [1, n] \longrightarrow \mathbb{R}^+$

$$-\sigma \in \Sigma$$

- every position
$$i, i \in [1, n]$$

 $W(g, \sigma, i) = \sum_{\{j \mid 1 \leq j \leq i-1 \land x_j = \sigma\}} g(j).$

i		1	2	3	4	5	6	7	8	9	10	11	12	
Т		d	b	с	a	b	с	b	с	a	a	a	a	_
	g(i)	4	1	3	1	3	2	1	6	1	2	1	3	_
	$\overline{W(g,\sigma,i)} = \sum_{\{j \mid 1 \leq j \leq 10 \land x_j = b\}} g(j).$													

<ロト < 団 > < 巨 > < 巨 > < 巨 < 三 の < ○</p>

15 / 34

Selective Weighted Coding

Experimental Results

Backward Weighted Coding

$W(g, \sigma, i)$ definition:

Given T = T[1, n] over Σ , define a general weight $W(g, \sigma, i)$

$$-g:[1,n]\longrightarrow {\rm I\!R}^+$$

$$-\sigma\in\Sigma$$

- every position
$$i, i \in [1, n]$$

 $W(g, \sigma, i) = \sum_{\{j \mid 1 \leq j \leq i-1 \land x_j = \sigma\}} g(j).$

i		1	2	3	4	5	6	7	8	9	10	11	12	
<u>T</u>		d	b	с	a	b	с	b	с	a	a	a	a	_
	g(i)	4	1	3	1	3	2	1	6	1	2	1	3	
	$W(g,\sigma,i) = \sum_{\{j \mid 1 \leq j \leq 10 \land x_j = b\}} g(j).$													

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □

16/34

Backward Weighted Coding b-adp and b-2

b-adp: A classic backward coding,

$$g=\mathbb{1}\equiv g(i)=1$$

b-2: Divides all the frequencies by 2 at the end of every block of k characters, for a given parameter k.

$$g_{b-2}(i+k) = 2g_{b-2}(i),$$

for each pair of indices i and i + k.

<ロ > < 回 > < 巨 > < 三 > < 三 > 三 の < で 17/34

Selective Weighted Coding

Experimental Results

Backward Weighted Coding b-w

b-w: Provides a fixed ratio of 2 between blocks but with rather smoother differences at the block borders.

$$g_{\text{b-w}}(i) = \left(\sqrt[k]{2}\right)^{i-1},$$

for $i \ge 1$, for a given parameter k.

Selective Weighted Coding

b-adp *vs.* b-w

i		1	2	3	4	5	6	7	8	9	10	11	12
т_		d	b	с	a	b	с	b	с	a	a	a	a
	g(i)	1	1	1	1	1	1	1	1	1	1	1	1
b-adp	W	1	1	1	1	2	2	3	3	2	3	4	5
_	IC	2	2.322	2.585	2.807	2	2.170	1.737	1.874	2.585	2.115	1.807	1.585
b-w	g(i)	1	1.414	2	2.828	4	5.657	8	11.314	16	22.627	32	45.255
	W	1	1	1	1	2.414	3	6.414	8.657	3.828	19.828	42.456	74.456
_	IC	2	2.322	2.681	3.073	2.219	2.345	1.704	1.739	3.393	1.503	0.893	0.574

The sum of the IC values: b-adp 25.588 bit b-w 24.447 bit.

Outline

Introduction

- Weighted Coding
- Selective Weighted Coding
- **4** Experimental Results

Constant function f

Periodic selection process, which is controlled by a skip-function f.

$$f(s) = c$$
, where s is the number of updates so far.

Two different strategies.

- (a) Complete-selective: the model is updated every f(s) characters.
- (b) Subset-selective: the model is based on the distributions of characters appearing at positions according to f(s).

5

6

7

Selective Weighted Coding

COMPLETE-selective
$$(T = x_1 \cdots x_n, g, f)$$

- 1 $s \leftarrow 0$; *last* $\leftarrow 0$; Initialize the model according to the uniform distribution on Σ
- 2 for $i \leftarrow 1$ to n do
- encode x_i according to the current model 3

4 if
$$i - last = f(s)$$
 then

update the model according to the distribution of the characters in Σ , given by the probabilities $\left\{ W(g,\sigma,i+1)/CW[1,i+1] \right\}_{\sigma \in \Sigma}$ $s \leftarrow s + 1$ – i

5

6

7

8

9

Weighted Coding 0000000

SUBSET-selective (
$$T = x_1 \cdots x_n, g, f$$
)

- 1 $s \leftarrow 0; \textit{ last} \leftarrow 0$; Initialize the model according to the uniform distribution on Σ
- ² for $i \leftarrow 1$ to n do
- a encode x_i according to the current model

4 if
$$i - last = f(s)$$
 then

update the model according to the distribution of the characters in Σ , given by the probabilities $\{W(g, \sigma, i+1)/CW[1, i+1]\}_{\sigma \in \Sigma}$ $s \leftarrow s+1$ $last \leftarrow i$ else $g(i) \leftarrow 0$

Selective Weighted Coding

Experimental Results

Complete Selective with s = 3

i		1	2	3	4	5	6	7	8	9	10	11	12
T.		d	b	с	a	b	с	b	с	a	a	a	a
b-adp	g(i)	1	1	1	1	1	1	1	1	1	1	1	1
	W	1	1	1	1	2	2	3	3	2	3	3	3
	IC	2	2	2	2.807	1.807	1.807	1.737	1.737	2.322	2.115	2.115	2.115
b-w	g(i)	1	1.414	2	2.828	4	5.657	8	11.314	16	22.627	32	45.255
	W	1	1	1	1	2.414	3	6.414	8.657	3.828	19.828	19.828	19.828
	IC	2	2	2	3.073	1.801	1.488	1.704	1.272	2.449	1.503	1.503	1.503

Sum of the IC values: b-adp is 24.564 bits b-w is 22.296 bits.

・ロト・日本・山本・山本・山本・

24 / 34

Neighted Codinı 2000000 Selective Weighted Coding

Experimental Results

Subset Selective with s = 3

i		1	2	3	4	5	6	7	8	9	10	11	12
т_		d	b	с	a	b	с	b	с	a	a	a	a
	g(i)	0	0	1	0	0	1	0	0	1	0	0	1
b-adp	W	1	1	1	1	1	2	1	3	1	2	2	2
	IC	2	2	2	2.322	2.322	1.322	2.585	1	2.585	1.807	1.807	1.807
	g(i)	0	0	2	0	0	5.657	0	0	16	0	0	45.255
b-w	W	1	1	1	1	1	3	1	8.657	1	17	17	17
	IC	2	2	2	2.585	2.585	1	3.543	0.429	3.543	0.702	0.702	0.702

Sum of IC values: b-adp - 23.558 b-w - 21.792.

<ロト < 回ト < 巨ト < 巨ト - 巨 - のへの

Varying function f

Opposing requirements:

- Weighted gives higher priority to positions close to the one currently processed.
- A need to attain as soon as possible a critical mass of selected items.

Neighted Codin 2000000 Experimental Results

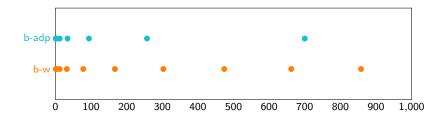
Tuned Selective Method

TUNED-SELECTIVE (
$$T = x_1 \cdots x_n$$
, g , threshold)
1 $cum \leftarrow 0$
2 for $i \leftarrow 1$ to n do
3 encode x_i according to the current model
4 $cum \leftarrow cum + g(i)/(g(1) + \cdots + g(i - 1))$
5 if $cum \ge threshold$ then
6 update the model
7 $cum \leftarrow 0$

Experimental Results 00000

Tuned Selective Method

decreasing number of updates for b-adp \Longleftrightarrow fixed intervals for b-w



threshold = 1.

<ロト < 団ト < 巨ト < 巨ト < 巨ト 三 の Q (C 28 / 34 Selective Weighted Coding 00000000●

Experimental Results

Tuned Selective Method

f(j) – the distance from the *j*-th selected location to the following one.

f(j) =round (j^{α})

α	indices of selected points																
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0.25	1	2	3	4	5	7	9	11	13	15	17	19	21	23	25	27	29
0.5	1	2	4	6	8	10	13	16	19	22	25	28	32	36	40	44	48
0.75	1	3	5	8	11	15	19	24	29	35	41	47	54	61	69	77	85
1	1	3	6	10	15	21	28	36	45	55	66	78	91	105	120	1 36	153
1.25	1	3	7	13	20	29	40	53	69	87	107	129	154	181	211	243	278
1.5	1	4	9	17	28	43	62	85	112	144	180	222	269	321	379	443	513

Sample of selected indices for various values of α .

29 / 34

Outline

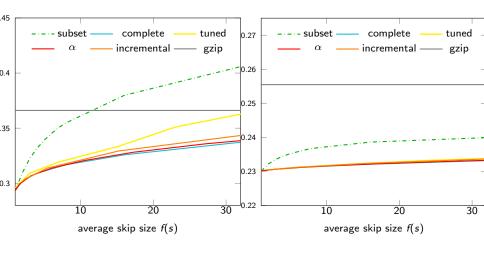
Introduction

- Weighted Coding
- Selective Weighted Coding
- Experimental Results

Selective Weighted Coding

Experimental Results $0 \bullet 000$

Compression Efficiency

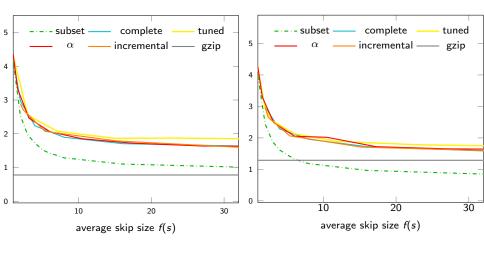


<ロト < 回 > < 巨 > < 巨 > < 巨 > 三 の < © 31/34

Selective Weighted Coding

Experimental Results

Encoding Time



<ロト < 回 > < 巨 > < 巨 > < 巨 > 三 の < © 32/34

Selective Weighted Coding

Experimental Results

Decoding Time

